Weather Scenario Generator: System Design and Project Status

Steven J. Lowe
SAIC
114 72nd St.
Virginia Beach, VA 23451
(757) 425-0052
steven.j.lowe@cpmx.saic.com

Eric A. Kihn
NGDC
325 Broadway E/GC2
Boulder, CO 80303
(303) 497-6346
eak@ngdc.noaa.gov

Ronald E. Englebretson
SAIC
550 Camino El Estero, Suite 205
Monterey, CA 93940
(408) 649-5242
englebretson@cpva.saic.com

Dr. Richard A. Siquig
Naval Research Laboratory
7 Grace Hopper Ave., MS 2
Monterey, CA 93943-5502
(408) 656-4732
siquig@nrlmry.navy.mil

Keywords
Synthetic Natural Environment, Weather, Scenario, Distributed Resources, Just-in-Time Production

ABSTRACT: *The Weather Scenario Generator (WSG) is intended to provide an integrated, physically consistent environmental data set meeting a Modeling and Simulation (M&S) customer's requirements for an authoritative and realistic representation of atmospheric, oceanic, and/or space natural environment elements for specified regions, time frames, and conditions. The WSG will automate the generation of a weather scenario, by allowing users to search for specific weather conditions as well as perform integrated searches of the Master Environmental Library (MEL) based on the functional relationships between the various data sources and online modeling capabilities. The WSG project is an FY-98 start under the DoD Modeling and Simulation Executive Agent for Air and Space Natural Environment. This paper will provide an overview of the design and functionality of the WSG system and its relationship to the MEL.*

1. Introduction

The Department of Defense (DoD) Modeling and Simulation (M&S) Master Plan[1] states that the next generation of M&S programs will require the inclusion of an integrated authoritative representation of the natural environment. The capability exists today to generate highly realistic weather scenarios on a wide range of scales, but the ability to properly define the scenario to meet the requirements of a given simulation has not yet been developed. The Weather Scenario Generator (WSG) project will develop a prototype system architecture aimed at meeting the 96 hour goal for the specification and development of a physically consistent, integrated synthetic natural environment (SNE) scenario.

The two primary goals for the project are:

1) Provide an intelligent search mechanism to locate desired conditions in historical archives
2) Demonstrate a just-in-time production capability to provide customized, realistic scenarios.

The WSG project has been funded by the M&S Executive Agent for Air and Space Natural Environment. The project is expected to meet the first goal above in FY-98, and demonstrate the second goal in mid FY-99. Because of the accelerated project schedule, the initial WSG prototype will focus on the generation of atmospheric scenarios only. To this end, the historical archive to be used is the NCEP/NCAR 40-year Reanalysis database, (described in Section 4), that provides global coverage of standard meteorological parameters at 6-hour resolution for 40 years. The just-in-time production capability will be supported by the Navy's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) that provides standard meteorological parameters at up to 1 km and 1 hr resolution. Even though the WSG project will focus on these two specific resources initially, the architecture will support the searching of any historical or climatological database, and the use of any model for just-in-time production. This flexibility will allow the WSG to be easily expanded to support ocean or space weather scenarios.

The area of scenario specification will also be kept extremely flexible, such that it can be customized to meet the needs of different communities. It is the eventual goal of the project to provide an interface to the WSG that allows the user to specify required parameters in their own terminology and to specify conditions in terms of the desired environmental effects on their simulated systems. However, the initial prototype will require that the user be able to specify the parameters and conditions desired in meteorological terms.

This paper provides an overview of the WSG system design and functionality. In Section 2, the concept of operations for the WSG is presented, including a description of the scenario development process and just-in-time production. Section 3 provides a detailed description of the WSG architecture and its planned interaction with the Master Environmental Library (MEL). Section 4 provides details on how the WSG project will implement and exploit the NCEP/NCAR 40-year Reanalysis database.

2. Concept of Operations

2.1 Overview

The concept of operations (CONOPS) for the WSG begins with the specification of the requirements for a weather scenario to support a simulation, including three basic classifications:

- The Spatial and Temporal Domain
- The List of Parameters Required
- The Types of Conditions Desired

The WSG will search a historical archive to find a date that provides the conditions desired in the region and season of interest. For the date range selected for development, the MEL will be queried to determine the best available resources to meet the specific requirements of the user. In addition, the WSG will have access to a just-in-time modeling capability that can provide a realistic depiction of the weather event selected at high spatial and temporal resolution. The final scenario will be developed by integrating the various resources of the MEL into one physically consistent representation of the atmosphere. The goal of the WSG is to deliver the final integrated scenario within 96 hours of the initial request.

2.2 Scenario Development

The specific sequence of events involved in developing a customized synthetic natural environment scenario using the WSG is as follows:

1) Scenario Specification
The staff METOC officer (SMO) for the simulation will work through the WSG desktop client application to specify the spatial and/or temporal restraints on the scenario, the parameters required in the scenario, and the conditions desired in the scenario.

2) Historical Search
The specified scenario will be submitted to the WSG Server for processing. The first step in this processing will be to find one or more dates in the NCEP/NCAR 40-year archive that are conducive to the conditions requested in the region specified. The list of dates will be sorted based on how well each date meets the desired set of conditions, and then returned to the WSG Client for presentation to the user.

3) Scenario Browse
The SMO will now have the option of browsing the conditions depicted in the NCEP/NCAR 40-yr database for any of the dates returned. The browse graphics will be

created in real time such that each user can create their own customized plots for the parameters of interest.

4) MEL Query
Once a date is selected for scenario development, multiple MEL queries will be performed to determine which MEL resources best meet the needs of the user. This will be determined based on how well each resource matches the spatial and temporal requirements of the scenario, as well as how many of the required parameters it provides.

5) Scenario Order
The MEL Query results will be presented to the SMO with an indication of how well each resource meets their scenario requirements. The SMO will then select the combination of resources that best meets their needs and begin the scenario ordering process. The ordering process will vary depending upon the number and complexity of the resources desired. The WSG Client application will integrate the MEL order process to the greatest extent possible to minimize the effort required by the SMO.

6) Scenario Delivery
The WSG Server will coordinate the ordering of all of the requested MEL resources. Once all of the orders are processed by MEL, and delivered to the WSG Server site, the SMO will be notified by email. He can than browse the final scenario via the WSG Client to insure that the scenario content is as desired. Once the SMO is satisfied with the content, he can request post-processing options such as re-gridding and/or re-formatting prior to final delivery.

2.3 Just-in-Time Production

The steps described above involve the specification and ordering of a scenario based on a weather event found in a historical archive. The MEL resources will always be checked first for possible data sources to support the desired scenario. But for cases where the highest level of fidelity and physical consistency is required, the WSG will also contain a just-in-time production site that will be capable of providing a highly realistic depiction of the selected weather event. Any model implemented at the WSG just-in-time production site will be registered in MEL such that it will be included as a result for any MEL query for which it is appropriate.

The Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS)[2] will be the first model utilized at the just-in-time production site. A request for a COAMPS generated scenario will be ordered through the MEL interface, and customized to generate output for the specific spatial and temporal domain required. The COAMPS will be initialized using the same NCEP/NCAR 40-year Reanalysis fields used for the historical search; thereby providing the most consistency between the COAMPS scenario and the historical event selected by the user.

3. System Architecture

3.1 Overview

The WSG will consist of a customizable client application that communicates with a server machine that will coordinate the searching, browsing, and production of scenarios. The WSG will also include a connection to the MEL, via the MEL Services API, to utilize the wide range of resources available through the MEL.

The complete WSG architecture, including its interactions with MEL, is shown in Figure 1. The components shown within the dashed blue square are planned to be physically located at the National Geophysical Data Center (NGDC) in Boulder, CO. The components shown within the dashed red square represent potential MEL resources.

The WSG Client will be a Java application that runs locally on the user's machine, *external* to a web browser. This avoids the possible security issues associated with Java applets and still allows for the reuse of many powerful Java classes that have already been developed. This Java application will be downloaded once from the WSG web site and establish its own direct TCP/IP connection with the WSG Server when run locally.

The WSG Server will be located at NGDC and coordinate all scenario development activities. It's main components are:

1. A web server
2. A communications layer
3. A search interface to the NCEP/NCAR database
4. An IDL plot server
5. An interface to the MEL Services API
6. An interface to a Just-in-Time modeling site

The web server will provide WSG project information as well as allow new users to download the WSG Client application. The communications layer will manage the individual client connections to the server, and coordinate communication with the other modules making up the server.

The following sections describe how each of the scenario development steps described in Section 2.2 will be implemented in the WSG Client-Server architecture.

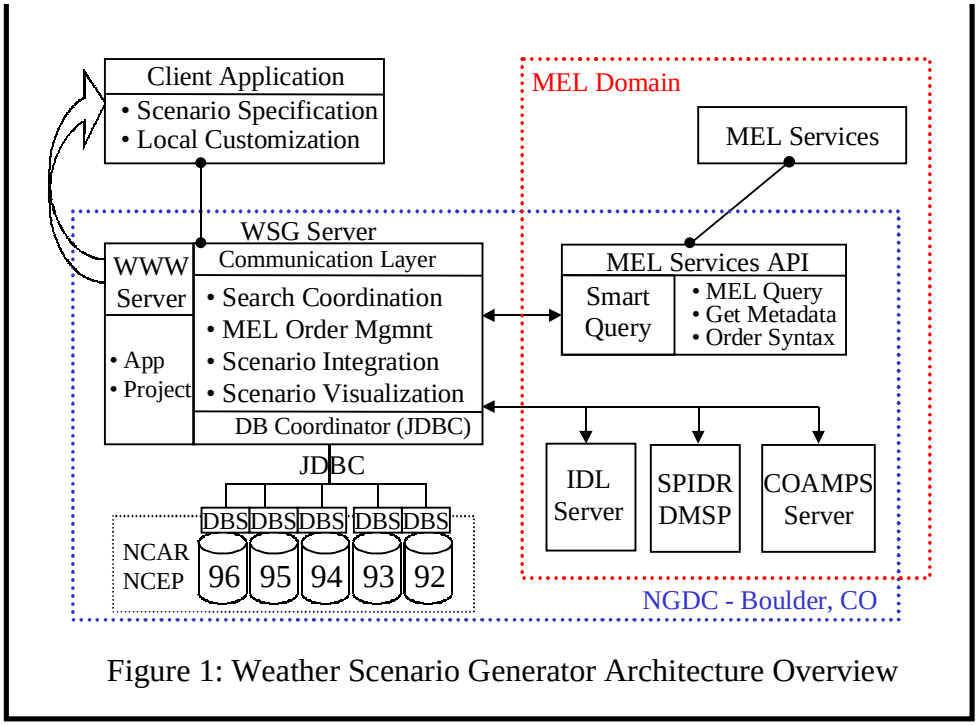

Figure 1: Weather Scenario Generator Architecture Overview

3.2 Scenario Specification

The Scenario Specification portion of the WSG Client application will be very similar to the existing MEL Query applet, with individual panels for: the spatial domain; the temporal domain; the conditions desired; and the list of parameters required. Each of these are discussed in further detail below.

Spatial Domain:
As with the MEL Query applet, the user will specify a region of interest by either directly inputting the latitude/longitude coordinates, or by dragging a bounding box on the map. Once the region is specified, the spatial display will be zoomed and dots representing the NCEP/NCAR database grid point locations will be displayed. The user will be prompted to select the points within the region on which to enforce the specified condition requirements. This information will allow the search module at the WSG Server to weight the dates found by how well they match the conditions at the individual points selected.

Temporal Domain:
The temporal domain will be specified by a contiguous range of months to be searched in the NCEP/NCAR database. The same set of months will be used for each year available for searching. This will allow the user to restrict the search to a desired month or season, or to leave the temporal domain unrestricted. In addition, the user will be able to simply specify a date range of interest to be extracted from the NCEP/NCAR database for browsing.

Conditions Desired:
The conditions can be specified in a number of ways, depending on the user's familiarity with the region of interest. The expert user could simply list the parameters of interest in determining the conditions (ie cloud cover, wind) and the WSG Server will return timelines of these parameters for the date range and position specified. The user could also specify min/max ranges on one or more of the parameters supported for searching. Conditions can also be specified via abstract definitions for each parameter. For instance, temperature conditions could be specified as "hot", "cold", or "typical". The various condition searching schemes are discussed in more detail in Section 4.3.

The ability to define abstract weather events will also be supported using simple rules-based logic built into the WSG Client application. For instance, a "winter storm" can be defined by setting the pressure to "low", the temperature to "cold", and the precipitation to "heavy". The WSG Client will contain a number of pre-defined weather events that, when selected, simply set the choices in the conditions panel to the required settings to locate the desired event. This approach will also allow the user to easily review the WSG interpretation of the desired weather event, and change the definition as required.

Parameters Required:
The list of parameters required for the scenario will be chosen from a database of parameter keywords maintained locally as part of the WSG Client application. This database of keywords will be mapped to the appropriate MEL keyword database at the WSG Server site prior to submitting queries via the MEL Services API. The advantage of a locally maintained keyword database is that users can customize the WSG Client application to use their own terminology.

3.3 Historical Searching

The result of the Scenario Specification module of the client application will be a formatted message sent to the WSG Server for processing. This message will contain the following components:

1. The Latitude/Longitude point(s) to search
2. The Month(s) to search
3. Flag indicating the query type
4. One or more "condition" strings to process
5. One or more parameters required

The WSG server will then interact with the NCEP/NCAR database via the intelligent search module. This search module will support multiple query modes, but in each case will result in a list of dates sorted by how well they satisfy the user's scenario requirements. The list of dates is returned to the client application for review by the user.

It is important to note that this intelligent search module will be developed such that it could be used for querying any database with a Java Data Base Compliant (JDBC) interface. This will allow the intelligent searching of other historical archives in addition to the NCEP/NCAR database.

3.4 Scenario Browsing

The list of dates determined by the intelligent search module is returned to the client application for presentation to the user. For each date, the user can generate customized plots for any parameter of interest. These plots will be generated at interactive speeds so that the user can effectively explore each proposed weather event to whatever level of detail they feel is required.

The plots used for scenario browsing will be generated using an Interactive Data Language (IDL) [3] plot server. The WSG Server will coordinate the flow of data from the NCEP/NCAR database to the IDL server. IDL provides a graphical user interface (GUI) as a Java class that will be integrated into the WSG Client application to allow the user to specify customized graphics.

In addition to the IDL plot server, IR and Visible satellite imagery will be available from the Defense Meteorological Satellite Program (DMSP) archives contained in the Space Physics Interactive Data Resource (SPIDR) [4] system at NGDC. DMSP imagery is very useful to meteorologists in visualizing and understanding weather events.

3.5 MEL Query

The Master Environmental Library (MEL) provides access to a wide array of geospatial data, products, and services, which are discovered by specifying a query consisting of a spatial and temporal domain, and an optional list of keywords. Once the WSG user has selected the weather event that best satisfies the requirements of their simulation, the WSG Server has all of the information required to perform a MEL query. The WSG Server will use the selected date range and the previously specified spatial domain and parameter list to query the MEL for appropriate resources. The implementation of the MEL Services API in FY-98 will allow applications such as WSG to perform MEL queries via a direct connection to the MEL Access Site.

The WSG Server will contain a customized "Smart Query" function that provides a more useful searching mechanism for the MEL. The current MEL query function is unable to sort out which metadata records were selected for each keyword in a multiple-keyword query. For instance, if a user queries MEL with keywords "A, B, C, D" and gets back 10 metadata records, there is no way to quickly (or automatically) determine which, if any, metadata records contain keyword "B". The ability to sort the metadata records by how many keywords they matched and to determine which metadata records contain which keywords is required. The Smart Query function will process multiple-keyword queries by breaking them up into multiple single-keyword queries and then employ set logic on the results of each query to determine which metadata records provide the best match based on the number of keywords provided.

3.6 Scenario Ordering

The process of ordering the final scenario will be streamlined by providing a single ordering interface built into the WSG Client application. The ordering interface will verify all of the scenario requirements, including the region and date to be ordered, the spatial and temporal resolution, the complete parameter list, and the set of

MEL resources to be ordered. The WSG Server will contain a scenario ordering module that will process all of the above information and submit individual MEL orders for the requested resources.

All MEL orders will be delivered to the WSG Server site unless the user explicitly requests direct delivery of the data sets to their location. The scenario ordering module will track the progress of each of the individual orders and notify the WSG user by email when all orders have been delivered.

3.7 Post-Processing and Delivery

Once all of the individual scenario components have been delivered to the WSG Server site, the user will have a number of post-processing options available. The ability to browse the final scenario will be provided by the same IDL server described in Section 3.4. Although it will not be developed in the first phase of the project, the ability to blend the scenario components into one physically consistent, integrated scenario will be added as a post-processing option. Finally, the user will select from available delivery formats, as is appropriate for the data types contained in their scenario. The eventual goal of the WSG project is to deliver all scenarios in the Synthetic Environment Data Representation and Interchange Specification (SEDRIS) [5] format.

4.0 NCEP/NCAR 40-yr Reanalysis Database

4.1 Overview

The National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) Reanalysis Project [6], started in 1991, was motivated by the apparent "climate changes" that resulted from many changes introduced in the NCEP operational Global Data Assimilation System (GDAS). The goal of the project is to provide a consistent historical archive of the earth's atmosphere by applying a "frozen" state-of-the-art data assimilation system to the most extensive observational data set available.

The result of this project is the NCEP/NCAR 40-yr Reanalysis Database, which covers the period 1957 through 1996 with global 2.5 deg. fields at 6 hour resolution. The global atmospheric model used at the core of the data assimilation system is identical to the NCEP global model operationally implemented on 10 January 1995, except that the reanalysis version has less horizontal resolution. A complete discussion of the NCEP/NCAR 40-yr Reanalysis Database is contained in Reference 6.

The output fields from the reanalysis runs are classified into four distinct, yet somewhat subjective, levels of accuracy. They are defined as follows:

- **A:** The analysis variable is strongly influenced by observed data, and hence it is in the most reliable class.
- **B:** The analysis variable is directly affected by observational data, but the model also has a very strong influence.
- **C:** The analysis variable is derived solely from the model fields, forced by the data assimilation to remain close to the real atmosphere.
- **D:** The analysis variable is obtained directly from climatological values, and does not depend on the model.

A number of comparisons have been performed between the NCEP/NCAR 40-yr reanalysis database and various climatologies, observed data sets, and other reanalysis databases. There are significant differences noted in some cases, particularly in the class C fields predicted by the reanalysis model. However, it is difficult to determine which data source is the most accurate due to the vast array of quality control and modeling techniques used. The NCEP/NCAR database was selected for use in the development of the WSG prototype because it provides a realistic representation of the atmosphere at 6 hour intervals for the entire globe. The scientific validity of the database may be debated, but of more importance to the modeling and simulation community is that a realistic representation of any region and/or season of interest is available for use.

4.2 WSG Implementation of Database

The NCEP/NCAR 40-year Reanalysis Database contains a wide variety of output parameters in various grid and data formats. The WSG implementation of the NCEP/NCAR database will involve the use of the following parameters, resulting in approximately 3 GB of data per year:

Class A:
Air Temperature at 2 m, and 850, 500, and 250 mb levels
Geopotential Height at 850, 500, and 250 mb levels
U-wind at 10 m and 850, 500, and 250 mb levels
V-wind at 10 m and 850, 500, and 250 mb levels
Mean Sea Level Pressure
Surface Temperature

Class B:
Pressure Vertical Velocity at 850, 500, and 250 mb levels

Relative Humidity at surface, 850, 500, and 250 mb levels
Surface Lifted Index
Precipitable water

Class C:
Visible beam downward solar flux at the surface
Visible diffuse downward solar flux at the surface
Cloud Cover Percentage (hi-mid-low)
Cloud Tops & Bottoms (hi-mid-low)
Volumetric soil moisture (0-10 cm and 10-200 cm layers)
Temperature of the soil (0-10 cm and 10-200 cm layers)
Precipitation Rate at the surface
Convective Precip Rate at the surface
Water equiv of accum snow depth
Runoff

Class D:
Ice concentration (0 or 1)

The data will be loaded into an Empress database that is distributed across multiple machines, each machine managing one year of data. The database will be structured into Latitude-Month tables, and indexed on Longitude to optimize the searching of conditions at a point. With this structure, each table will be only 3 megabytes in size, which should allow for interactive query speeds.

For the initial WSG prototype work, only the most recent 5 years of the NCEP/NCAR database will be available for searching. The full 40-year archive will be added later.

4.3 Querying the Database

There will be at least three ways that the WSG will be able to query the NCEP/NCAR database. The method selected will depend upon the user's meteorological knowledge and experience in the region of interest. These query methods are outlined below.

General Browse:
An expert user of the system may want to just extract data for a specific region and time frame such that they can visually browse through conditions. For this user, the WSG system will allow the user to interactively plot the parameters of choice for any region and time frame desired.

Specific Browse:
Some users may have exact thresholds and/or limitations that must be maintained on certain parameters. For this user, the WSG will search the database using the spatial and temporal restraints set by the user for dates that best match the desired conditions. A sorted list of dates will be presented to the user, at which point the user can begin interactively plotting parameters of interest to determine which date they feel best represents the type of weather conditions desired.

Abstract Browse:
Because some users may not know or have exact requirements for their weather scenario, abstract searches will also be supported on a subset of the parameters in the database. For each parameter, three levels will be defined in abstract terms representing a high, typical, and low condition for that parameter. For example, the wind levels will be "windy", "typical", and "calm"; the temperature levels will be "hot", "typical", and "cold". The parameters involved in an abstract query will be extracted from the database for the region and timeframe specified. For each parameter, the dates will be sorted by value and the top, middle, and bottom five dates will be selected as the high, typical, and low condition, respectively. Finally, set logic will be use to sort the dates by how many conditions an individual date meets.

5. Conclusions

The Weather Scenario Generator (WSG) is intended to provide support to simulation developers seeking to include realistic weather as an integrated part of their simulated environment. The result of the WSG will be a physically consistent, integrated scenario that is delivered to the simulation developer in a standard format on a user-defined grid system. It will not address a dial-the-weather capability, but rather will seek to generate an integrated scenario based upon a real weather event from history that provided the desired conditions in the region of interest. In addition, the WSG does not address the computation and distribution of environmental effects at simulation runtime, but rather is intended to develop the underlying physically consistent synthetic environment that is required to maintain correlation among those computed effects.

6. References

[1] Department of Defense, Under Secretary of Defense for Acquisition and Technology. "Modeling and Simulation (M&S) Master Plan", October 1995. www.dmso.mil/docslib/mspolicy/msmp/1095msmp.

[2] R. M. Hodur: "The U.S. Navy's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS)" Monthly Weather Review, Vol. 125, pp. 1414-1430, July 1997.

[3] Interactive Data Language (IDL): Information available from Research Systems Inc. at www.rsinc.com or info@rsinc.com.

[4] E. Kihn: “Space Physics Interactive Data Repository (SPIDR) on the Web” Earth System Monitor, Vol. 6, No. 3, March 1996. Information available at www.ngdc.noaa.gov:8080.
[5] Synthetic Environment Data Representation and Interchange Specification (SEDRIS). Information available at www.sedris.org.
[6] E. Kalnay, R. Jenne et al.: “The NCEP/NCAR 40-Year Reanalysis Project” Bulletin of the American Meteorological Society, Vol. 77, No. 3, pp. 437-471, March 1996.

Author Biographies

STEVE LOWE is an atmospheric scientist for SAIC. He is currently the technical lead for the Weather Scenario Generator project and a member of the Master Environmental Library (MEL) development team. Mr. Lowe has been responsible for the generation and exchange of physically consistent, integrated synthetic natural environment scenarios within the MEL, and has lead the development of distributed computing techniques and GRIB applications for MEL. Mr. Lowe holds BS and MS degrees in Aerospace Engineering from Virginia Tech.

ERIC KIHN is a physicist with the National Geophysical Data Center in Boulder, CO. He is the lead developer on the Space Physics Interactive Data Repository (SPIDR) project, for which he was won Department of Commerce Silver Medals in 1994 and 1997. His research interests include space weather and application development involving relational databases and online interfaces. Mr. Kihn has a BA in Physics and an MS in Mathematics, both from the University of Colorado.

RON ENGLEBRETSON is a research meteorologist for SAIC. His research interests include: diagnosis and verification of global and mesoscale NWP model products; development of regional forecaster handbooks with emphasis on the correlation of NWP products, satellite data, and local observations to regional and local weather patterns; evaluations of ports used by navy vessels as tropical cyclone havens; and weather satellite imagery interpretation guides. For the past four years he has served as the Division Manager for the Monterey office of SAIC, whose primary customer is the Marine Meteorology Division of the Naval Research Laboratory. Mr. Englebretson received his formal education, MS in Meteorology, from the Naval Postgraduate school, and his practical education from his combined 25 years as a navy meteorologist and 17 years as a SAIC research meteorologist.

DR. RICHARD SIQUIG is a physicist with the Marine Meteorology Division of the Naval Research Laboratory. He is the project manager for the development and fielding of the Master Environmental Library (MEL). He has conducted prior research in environmental effects on unmanned aerial vehicles and the determination of ship tracks in clouds and is currently leading the Weather Scenario Generator Project. Dr. Siquig holds a Ph.D. in Astrophysics from the University of Colorado.